



OIPE

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/031,496A

DATE: 05/13/2003

TIME: 16:19:32

Input Set : A:\PCT99-45.txt

Output Set: N:\CRF4\05132003\J031496A.raw

```

3 <110> APPLICANT: Adney, William S.
4      Decker, Stephen R.
5      Lantz-McCarter, Suzanne
6      Baker, John O.
7      Vinzant, Todd B.
8      Nieves, Rafael A.
9      Himmel, Michael E.
11 <120> TITLE OF INVENTION: CELLOBIOHYDROLASE REDUCED GLYCOSYLATION VARIANTS: CBHIN45A;
12      CBHIN270A; AND CBHIN384A
14 <130> FILE REFERENCE: HIMMEL NFEL IR# 99-45
C--> 16 <140> CURRENT APPLICATION NUMBER: US/10/031,496A
C--> 17 <141> CURRENT FILING DATE: 2003-05-06
19 <160> NUMBER OF SEQ ID NOS: 4
21 <170> SOFTWARE: PatentIn Ver. 2.0
23 <210> SEQ ID NO: 1
24 <211> LENGTH: 496
25 <212> TYPE: PRT
26 <213> ORGANISM: Trichoderma reesei
28 <400> SEQUENCE: 1
29 Gln Ser Ala Cys Thr Leu Gln Ser Glu Thr His Pro Pro Leu Thr Trp
30 1 5 10 15
32 Gln Lys Cys Ser Ser Gly Gly Thr Cys Thr Gln Gln Thr Gly Ser Val
33 20 25 30
35 Val Ile Asp Ala Asn Trp Arg Trp Thr His Ala Thr Ala Ser Ser Thr
36 35 40 45
38 Asn Cys Tyr Asp Gly Asn Thr Trp Ser Ser Thr Leu Cys Pro Asp Asn
39 50 55 60
41 Glu Thr Cys Ala Lys Asn Cys Cys Leu Asp Gly Ala Ala Tyr Ala Ser
42 65 70 75 80
44 Thr Tyr Gly Val Thr Thr Ser Gly Asn Ser Leu Ser Ile Gly Phe Val
45 85 90 95
47 Thr Gln Ser Ala Gln Lys Asn Val Gly Ala Arg Leu Tyr Leu Met Ala
48 100 105 110
50 Ser Asp Thr Thr Tyr Gln Glu Phe Thr Leu Leu Gly Asn Glu Phe Ser
51 115 120 125
53 Phe Asp Val Asp Val Ser Gln Leu Pro Cys Gly Leu Asn Gly Ala Leu
54 130 135 140
56 Tyr Phe Val Ser Met Asp Ala Asp Gly Gly Val Ser Lys Tyr Pro Thr
57 145 150 155 160
59 Asn Thr Ala Gly Ala Lys Tyr Gly Thr Gly Tyr Cys Asp Ser Gln Cys
60 165 170 175
62 Pro Arg Asp Leu Lys Phe Ile Asn Gly Gln Ala Asn Val Glu Gly Trp
63 180 185 190

```

## RAW SEQUENCE LISTING

DATE: 05/13/2003

PATENT APPLICATION: US/10/031,496A

TIME: 16:19:32

Input Set : A:\PCT99-45.txt

Output Set: N:\CRF4\05132003\J031496A.raw

```

65 Glu Pro Ser Ser Asn Asn Ala Asn Thr Gly Ile Gly Gly His Gly Ser
66      195      200      205
68 Cys Cys Ser Gln Met Asp Ile Trp Glu Ala Asn Ser Ile Ser Glu Ala
69      210      215      220
71 Leu Thr Pro His Pro Cys Thr Thr Val Gly Gln Glu Ile Cys Glu Gly
72      225      230      235      240
74 Asp Gly Cys Gly Gly Thr Tyr Ser Asp Asn Arg Tyr Gly Gly Thr Cys
75      245      250      255
77 Asp Pro Asp Gly Cys Asp Trp Asn Pro Tyr Arg Leu Gly Asn Thr Ser
78      260      265      270
80 Phe Tyr Gly Pro Gly Ser Ser Phe Thr Leu Asp Thr Thr Lys Lys Leu
81      275      280      285
83 Thr Val Val Thr Gln Phe Glu Thr Ser Gly Ala Ile Asn Arg Tyr Tyr
84      290      295      300
86 Val Gln Asn Gly Val Thr Phe Gln Gln Pro Asn Ala Glu Leu Gly Ser
87      305      310      315      320
89 Tyr Ser Gly Asn Glu Leu Asn Asp Asp Tyr Cys Thr Ala Glu Glu Ala
90      325      330      335
91 Glu Phe Gly Gly Ser Ser Phe Ser Asp Lys Gly Gly Leu Thr Gln Phe
92      340      345      350
94 Lys Lys Ala Thr Ser Gly Gly Met Val Leu Val Met Ser Leu Trp Asp
95      355      360      365
98 Asp Tyr Tyr Ala Asn Met Leu Trp Leu Asp Ser Thr Tyr Pro Thr Asn
99      370      375      380
101 Glu Thr Ser Ser Thr Pro Gly Ala Val Arg Gly Ser Cys Ser Thr Ser
102      385      390      395      400
104 Ser Gly Val Pro Ala Gln Val Glu Ser Gln Ser Pro Asn Ala Lys Val
105      405      410      415
107 Thr Phe Ser Asn Ile Lys Phe Gly Pro Ile Gly Ser Thr Gly Asn Pro
108      420      425      430
110 Ser Gly Gly Asn Pro Pro Gly Gly Asn Arg Gly Thr Thr Thr Arg
111      435      440      445
113 Arg Pro Ala Thr Thr Thr Gly Ser Ser Pro Gly Pro Thr Gln Ser His
114      450      455      460
116 Tyr Gly Gln Cys Gly Gly Ile Gly Tyr Ser Gly Pro Thr Val Cys Ala
117      465      470      475      480
119 Ser Gly Thr Thr Cys Gln Val Leu Asn Pro Tyr Tyr Ser Gln Cys Leu
120      485      490      495
123 <110> SEQ ID NO: 2
124 <211> LENGTH: 496
125 <212> TYPE: PRT
126 <213> ORGANISM: Trichoderma reesei
128 <400> SEQUENCE: 1
129 Gln Ser Ala Cys Thr Leu Gln Ser Glu Thr His Pro Pro Leu Thr Trp
130      1      5      10      15
132 Gln Lys Cys Ser Ser Gly Gly Thr Cys Thr Gln Gln Thr Gly Ser Val
133      20      25      30
135 Val Ile Asp Ala Asn Trp Arg Trp Thr His Ala Thr Asn Ser Ser Thr
136      35      40      45

```

## RAW SEQUENCE LISTING

DATE: 05/13/2003

PATENT APPLICATION: US/10/031,496A

TIME: 16:19:32

Input Set : A:\PCT99-45.txt

Output Set: N:\CRF4\05132003\J031496A.raw

```

138 Asn Cys Tyr Asp Gly Asn Thr Trp Ser Ser Thr Leu Cys Pro Asp Asn
139      50                      55                      60
141 Glu Thr Cys Ala Lys Asn Cys Cys Leu Asp Gly Ala Ala Tyr Ala Ser
142 65                      70                      75                      80
144 Thr Tyr Gly Val Thr Thr Ser Gly Asn Ser Leu Ser Ile Gly Phe Val
145      85                      90                      95
147 Thr Gln Ser Ala Gln Lys Asn Val Gly Ala Arg Leu Tyr Leu Met Ala
148      100                     105                     110
150 Ser Asp Thr Thr Tyr Gln Glu Phe Thr Leu Leu Gly Asn Glu Phe Ser
151      115                     120                     125
153 Phe Asp Val Asp Val Ser Gln Leu Pro Cys Gly Leu Asn Gly Ala Leu
154      130                     135                     140
156 Tyr Phe Val Ser Met Asp Ala Asp Gly Gly Val Ser Lys Tyr Pro Thr
157 145                     150                     155                     160
159 Asn Thr Ala Gly Ala Lys Tyr Gly Thr Gly Tyr Cys Asp Ser Gln Cys
160      165                     170                     175
162 Pro Arg Asp Leu Lys Phe Ile Asn Gly Gln Ala Asn Val Glu Gly Trp
163      180                     185                     190
165 Glu Pro Ser Ser Asn Asn Ala Asn Thr Gly Ile Gly Gly His Gly Ser
166      195                     200                     205
168 Cys Cys Ser Glu Met Asp Ile Trp Glu Ala Asn Ser Ile Ser Glu Ala
169      210                     215                     220
171 Leu Thr Pro His Pro Cys Thr Thr Val Gly Gln Glu Ile Cys Glu Gly
172 225                     230                     235                     240
174 Asp Gly Cys Gly Gly Thr Tyr Ser Asp Asn Arg Tyr Gly Gly Thr Cys
175      245                     250                     255
177 Asp Pro Asp Gly Cys Asp Trp Asn Pro Tyr Arg Leu Gly Ala Thr Ser
178      260                     265                     270
180 Phe Tyr Gly Pro Gly Ser Ser Phe Thr Leu Asp Thr Thr Lys Lys Leu
181      275                     280                     285
183 Thr Val Val Thr Gln Phe Glu Thr Ser Gly Ala Ile Asn Arg Tyr Tyr
184      290                     295                     300
186 Val Gln Asn Gly Val Thr Phe Gln Gln Pro Asn Ala Glu Leu Gly Ser
187 305                     310                     315                     320
189 Tyr Ser Gly Asn Glu Leu Asn Asp Asp Tyr Cys Thr Ala Glu Glu Ala
190      325                     330                     335
192 Glu Phe Gly Gly Ser Ser Phe Ser Asp Lys Gly Gly Leu Thr Gln Phe
193      340                     345                     350
195 Lys Lys Ala Thr Ser Gly Gly Met Val Leu Val Met Ser Leu Trp Asp
196      355                     360                     365
198 Asp Tyr Tyr Ala Asn Met Leu Trp Leu Asp Ser Thr Tyr Pro Thr Asn
199      370                     375                     380
201 Glu Thr Ser Ser Thr Pro Gly Ala Val Arg Gly Ser Cys Ser Thr Ser
202 385                     390                     395                     400
204 Ser Gly Val Pro Ala Gln Val Glu Ser Gln Ser Pro Asn Ala Lys Val
205      405                     410                     415
207 Thr Phe Ser Asn Ile Lys Phe Gly Pro Ile Gly Ser Thr Gly Asn Pro
208      420                     425                     430
210 Ser Gly Gly Asn Pro Pro Gly Gly Asn Arg Gly Thr Thr Thr Thr Arg

```

## RAW SEQUENCE LISTING

DATE: 05/13/2003

PATENT APPLICATION: US/10/031,496A

TIME: 15:19:32

Input Set : A:\PCT99-45.txt

Output Set: N:\CRF4\05132003\J031496A.raw

```

211          435          440          445
213 Arg Pro Ala Thr Thr Thr Gly Ser Ser Pro Gly Pro Thr Gln Ser His
214          450          455          460
216 Tyr Gly Gln Cys Gly Gly Ile Gly Tyr Ser Gly Pro Thr Val Cys Ala
217 465          470          475          480
219 Ser Gly Thr Thr Cys Gln Val Leu Asn Pro Tyr Tyr Ser Gln Cys Leu
220          485          490          495
223 <10> SEQ ID NO: 3
224 <11> LENGTH: 490
225 <12> TYPE: PRT
226 <13> ORGANISM: Trichoderma reesei
227 <400> SEQUENCE:
229 Gln Ser Ala Cys Thr Leu Gln Ser Glu Thr His Pro Pro Leu Thr Trp
230 1 5 10 15
232 Gln Lys Cys Ser Ser Gly Gly Thr Cys Thr Gln Gln Thr Gly Ser Val
233 20 25 30
235 Val Ile Asp Ala Asn Trp Arg Trp Thr His Ala Thr Asn Ser Ser Thr
236 35 40 45
238 Asn Cys Tyr Asp Gly Asn Thr Trp Ser Ser Thr Leu Cys Pro Asp Asn
239 50 55 60
241 Glu Thr Cys Ala Lys Asn Cys Cys Leu Asp Gly Ala Ala Tyr Ala Ser
242 65 70 75 80
244 Thr Tyr Gly Val Thr Thr Ser Gly Asn Ser Leu Ser Ile Gly Phe Val
245 85 90 95
247 Thr Gln Ser Ala Gln Lys Asn Val Gly Ala Arg Leu Tyr Leu Met Ala
248 100 105 110
250 Ser Asp Thr Thr Tyr Gln Glu Phe Thr Leu Leu Gly Asn Glu Phe Ser
251 115 120 125
253 Phe Asp Val Asp Val Ser Gln Leu Pro Cys Gly Leu Asn Gly Ala Leu
254 130 135 140
256 Tyr Phe Val Ser Met Asp Ala Asp Gly Gly Val Ser Lys Tyr Pro Thr
257 145 150 155 160
259 Asn Thr Ala Gly Ala Lys Tyr Gly Thr Gly Tyr Cys Asp Ser Gln Cys
260 165 170 175
262 Pro Arg Asp Leu Lys Phe Ile Asn Gly Gln Ala Asn Val Glu Gly Trp
263 180 185 190
265 Glu Pro Ser Ser Asn Asn Ala Asn Thr Gly Ile Gly Gly His Gly Ser
266 195 200 205
268 Cys Cys Ser Glu Met Asp Ile Trp Glu Ala Asn Ser Ile Ser Glu Ala
269 210 215 220
271 Leu Thr Pro His Pro Cys Thr Thr Val Gly Gln Glu Ile Cys Glu Gly
272 225 230 235 240
274 Asp Gly Cys Gly Gly Thr Tyr Ser Asp Asn Arg Tyr Gly Gly Thr Cys
275 245 250 255
277 Asp Pro Asp Gly Cys Asp Trp Asn Pro Tyr Arg Leu Gly Asn Thr Ser
278 260 265 270
280 Phe Tyr Gly Pro Gly Ser Ser Phe Thr Leu Asp Thr Thr Lys Lys Leu
281 275 280 285
283 Thr Val Val Thr Gln Phe Glu Thr Ser Gly Ala Ile Asn Arg Tyr Tyr

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/031,496A

DATE: 05/13/2003

TIME: 16:19:32

Input Set : A:\PCT99-45.txt

Output Set: N:\CRF4\05132003\J031496A.raw

```

284      290      295      300
286 Val Gln Asn Gly Val Thr Phe Gln Gln Pro Asn Ala Glu Leu Gly Ser
287 305      310      315      320
288 Tyr Ser Gly Asn Glu Leu Asn Asp Asp Tyr Cys Thr Ala Glu Glu Ala
289      325      330      335
291 Glu Phe Gly Gly Ser Ser Phe Ser Asp Lys Gly Gly Leu Thr Gln Phe
292      340      345      350
295 Lys Lys Ala Thr Ser Gly Gly Met Val Leu Val Met Ser Leu Trp Asp
296      355      360      365
298 Asp Tyr Tyr Ala Asn Met Leu Trp Leu Asp Ser Thr Tyr Pro Thr Ala
299      370      375      380
301 Glu Thr Ser Ser Thr Pro Gly Ala Val Arg Gly Ser Cys Ser Thr Ser
302 385      390      395      400
304 Ser Gly Val Pro Ala Gln Val Glu Ser Gln Ser Pro Asn Ala Lys Val
305      405      410      415
307 Thr Phe Ser Asn Ile Lys Phe Gly Pro Ile Gly Ser Thr Gly Asn Pro
308      420      425      430
310 Ser Gly Gly Asn Pro Pro Gly Gly Asn Arg Gly Thr Thr Thr Thr Arg
311      435      440      445
313 Arg Pro Ala Thr Thr Thr Gly Ser Ser Pro Gly Pro Thr Gln Ser His
314      450      455      460
316 Tyr Gly Gln Cys Gly Gly Ile Gly Tyr Ser Gly Pro Thr Val Cys Ala
317 465      470      475      480
319 Ser Gly Thr Thr Cys Gln Val Leu Asn Pro Tyr Tyr Ser Gln Cys Leu
320      485      490      495
323 <210> SEQ ID NO: 4
324 <211> LENGTH: 496
325 <212> TYPE: PRT
326 <213> ORGANISM: Trichoderma reesei
328 <400> SEQUENCE: 4
329 Gln Ser Ala Cys Thr Leu Gln Ser Glu Thr His Pro Pro Leu Thr Trp
330 1      5      10      15
332 Gln Lys Cys Ser Ser Gly Gly Thr Cys Thr Gln Gln Thr Gly Ser Val
333      20      25      30
335 Val Ile Asp Ala Asn Trp Arg Trp Thr His Ala Thr Asn Ser Ser Thr
336      35      40      45
338 Asn Cys Tyr Asp Gly Asn Thr Trp Ser Ser Thr Leu Cys Pro Asp Asn
339      50      55      60
341 Glu Thr Cys Ala Lys Asn Cys Cys Leu Asp Gly Ala Ala Tyr Ala Ser
342 65      70      75      80
344 Thr Tyr Gly Val Thr Thr Ser Gly Asn Ser Leu Ser Ile Gly Phe Val
345      85      90      95
347 Thr Gln Ser Ala Gln Lys Asn Val Gly Ala Arg Leu Tyr Leu Met Ala
348      100      105      110
350 Ser Asp Thr Thr Tyr Gln Glu Phe Thr Leu Leu Gly Asn Glu Phe Ser
351      115      120      125
353 Phe Asp Val Asp Val Ser Gln Leu Pro Cys Gly Leu Asn Gly Ala Leu
354      130      135      140
356 Tyr Phe Val Ser Met Asp Ala Asp Gly Gly Val Ser Lys Tyr Pro Thr

```

RAW SEQUENCE LISTING ERROR SUMMARY  
PATENT APPLICATION: US/10/031,496ADATE: 05/13/2003  
TIME: 10:19:33

Input Set : A:\PCT99-45.txt

Output Set: N:\CRF4\05132003\J031496A.raw

## Invalid Line Length:

The rules require that a line not exceed 72 characters in length. This includes spaces.

Seq#:1; Line(s) 2,3,4,5,6,7,8,9,10,13,14,15,16,17,18,19,20,21,22,23,24,25  
Seq#:1; Line(s) 26,27,28,29,30,31,32,33,34,35,36,37,38,39,40,41,42,43,44,45  
Seq#:1; Line(s) 46,47,48,49,50,51,52,53,54,55,56,57,58,59,60,61,62,63,64,65  
Seq#:1; Line(s) 66,67,68,69,70,71,72,73,74,75,76,77,78,79,80,81,82,83,84,85  
Seq#:1; Line(s) 86,87,88,89,90,91,92,93,94,95,96,97,98,99,100,101,102,103  
Seq#:1; Line(s) 104,105,106,107,108,109,110,111,112,113,114,115,116,117,118  
Seq#:1; Line(s) 119,120,121,122,123  
Seq#:2; Line(s) 124,125,126,127,128,129,130,131,132,133,134,135,136,137,138,139  
Seq#:2; Line(s) 140,141,142,143,144,145,146,147,148,149,150,151,152,153,154,155  
Seq#:2; Line(s) 156,157,158,159,160,161,162,163,164,165,166,167,168,169,170,171  
Seq#:2; Line(s) 172,173,174,175,176,177,178,179,180,181,182,183,184,185,186,187  
Seq#:2; Line(s) 188,189,190,191,192,193,194,195,196,197,198,199,200,201,202,203  
Seq#:2; Line(s) 204,205,206,207,208,209,210,211,212,213,214,215,216,217,218  
Seq#:2; Line(s) 219,220,221,222,223  
Seq#:3; Line(s) 224,225,226,227,228,229,230,231,232,233,234,235,236,237,238,239  
Seq#:3; Line(s) 240,241,242,243,244,245,246,247,248,249,250,251,252,253,254,255,256,257  
Seq#:3; Line(s) 258,259,260,261,262,263,264,265,266,267,268,269,270,271,272,273,274,275  
Seq#:3; Line(s) 276,277,278,279,280,281,282,283,284,285,286,287,288,289,290,291  
Seq#:3; Line(s) 292,293,294,295,296,297,298,299,300,301,302,303,304,305,306  
Seq#:3; Line(s) 307,308,309,310,311,312,313,314,315,316,317,318,319,320,321  
Seq#:3; Line(s) 322,323  
Seq#:4; Line(s) 324,325,326,327,328,329,330,331,332,333,334,335,336,337,338,339  
Seq#:4; Line(s) 340,341,342,343,344,345,346,347,348,349,350,351,352,353,354,355,356,357  
Seq#:4; Line(s) 358,359,360,361,362,363,364,365,366,367,368,369,370,371,372,373,374,375  
Seq#:4; Line(s) 376,377,378,379,380,381,382,383,384,385,386,387,388,389,390,391,392  
Seq#:4; Line(s) 393,394,395,396,397,398,399,400,401,402,403,404,405,406,407  
Seq#:4; Line(s) 408,409,410,411,412,413,414,415,416,417,418,419,420

## VERIFICATION SUMMARY

PATENT APPLICATION: US/10/031,496A

DATE: 05/13/2003

TIME: 16:19:33

Input Set : A:\PCT99-45.txt

Output Set: N:\CRF4\05132003\J031496A.raw

L:16 M:270 C: Current Application Number differs, Replaced Application Number  
L:17 M:271 C: Current Filing Date differs, Replaced Current Filing Date